ON THE IMPORTANCE OF CORRESPONDENCE BETWEEN SHAPES AND TIMBRE

Mohammad Adeli

Uinversité de Sherbrooke
NECOTIS Lab
Sherbrooke, Canada
mohammad.adeli@usherbrooke.ca

Jean Rouat

Uinversité de Sherbrooke
NECOTIS Lab
Sherbrooke, Canada
jean.rouat@usherbrooke.ca

Stéphane Molotchnikoff

Uinversité de Montréal
Neuroscience Lab
Montréal, Canada
stephane.molotchnikoff@umontreal.ca

ABSTRACT

The results of a preliminary study of the audio-visual correspondence between musical timbre and visual shapes are reported. 22 participants had to play 20 musical sounds and choose a shape for each. An association between timbre and visual shapes emerged. Soft timbres seem to match with rounded shapes, harsh timbres with sharp angular shapes and timbres having elements of softness and harshness together with a mixture of the two previous shapes. The correspondence between timbres and shapes should lend itself to the development of perceptually supported musical interfaces and substitution systems. A larger scale experiment with more sounds and participants is underway and confirms the preliminary results reported in this paper.

1. INTRODUCTION

For persons with vision-audition synesthesia, visual objects give rise to the involuntary audition of sounds or vice versa. Studies have shown that normal people can similarly match light intensity with auditory loudness [1] and auditory pitch [1, 2]; visual size and elevation with auditory pitch [3]. Such audio-visual correspondences have been used to develop interactive interfaces for production of audio-visual music [4], sound visualization [5], improvement of sound synthesis using sound visualizations [6]. However, in most studies simple stimuli (e.g. simple tones) have been used which do not permit to study more important perceptual features such as auditory timbre and visual shapes. For instance, non-musicians can recognize Piano from its timbre while they do not perceive pitch difference between two close notes. Similarly visual shapes play a key role in the identification of visual objects. Therefore, in this paper, potential associations between shapes and timbres are investigated and results of a preliminary experiment are reported.

2. EXPERIMENT

The experiment was conducted online. There were 20 sounds (five instruments by four pitches) which were presented in randomized order twice; once against colored shapes and once against grayscale shapes.

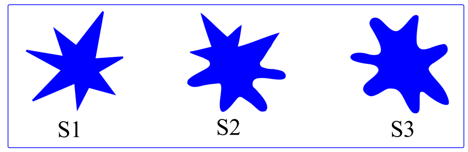

Figure 1: Shapes used as visual stimuli: each web page included 4 rows of shapes. One row of shapes is shown in the figure. The other three rows included the same shapes in the same order. Each row had a different color (grayscale) which varied randomly. In a given page of the web interface, shapes appeared in either 4 colors (blue, green, red and yellow) or 4 grayscales (black=0, 0.3, 0.6 and 0.9). Subjects played the sound and selected a shape for it.

Sounds of Electric Piano, Marimba, Classic Guitar, Cello, Tenor Sax each with four different pitches (100Hz, 150Hz, 200Hz, 250Hz) were normalized in energy and windowed so that they had the same envelope. Duration of each sound was one second. Three shapes were used as visual stimuli (see Figure 1). The leftmost (jagged) and the rightmost (rounded) shapes in Figure 1 have previously been used in Kiki-Bouba experiment [7]. The authors decided to use a mixture of these two shapes due to diversity of auditory timbres. On each web page, these shapes appeared either in 4 colors (blue, green, red and yellow) or in 4 grayscales (0 (black), 0.3, 0.6 and 0.9 (lightest scale)). Each row of shapes had a different color (or grayscale) but this color (or grayscale) varied randomly across pages. 22 subjects (5 females and 17 males) volontarily participated in the online experiment. Subjects had to play each sound and select a shape for it. No training was provided to the subjects.

3. RESULTS

Subjects have selected different shapes for different timbres. Table 1 summarizes the timbre-shape associations for both colored and gray scale shapes. Soft timbres (e.g. the sound derived from Piano) were associated with a rounded shape, harsh timbres (e.g. the sound derived from Cello at 250 Hz) with a sharp angular shape and sounds having elements of softness and harshness together (e.g. the sound derived from Sax) with a mixture of previous shapes. As seen in Table 1, the most selected shape for a given sound was not affected by color or grayscale. A larger scale experiment is still in progress [8].

Table 1: Shape selections: the percentage of subjects who selected each shape for a given sound. Symbols ✶ , ✪ and ✿ represent shapes S1, S2 and S3 in Figure 1. In general, soft timbres (e.g. windowed piano) have been associated with a rounded shape, harsh timbres (e.g. windowed Cello at 250 Hz) with a sharp angular shape and sounds having elements of harshness and softness together (e.g. windowed Saxophone) with a mixed shape. The most selected shape has not changed with grayscale or color. The chance level was 33%.

		Colored Shapes			Gray-scale shapes		
Timbre	Pitch	✶	✪	✿	✶	✪	✿
Cello	100Hz	0	**68.2**	31.8	4.6	**63.6**	31.8
	150Hz	18.2	27.3	**54.5**	13.6	**45.5**	40.9
	200Hz	40.9	**50**	9.1	22.7	**63.6**	13.7
	250Hz	**63.6**	31.8	4.6	**68.2**	18.2	13.6
Guitar	100Hz	0	**68.2**	31.8	0	**77.3**	22.7
	150Hz	4.5	41	**54.5**	4.6	31.8	**63.6**
	200Hz	22.7	**68.2**	9.1	22.7	**50**	27.3
	250Hz	31.8	**54.6**	13.6	18.2	**63.6**	18.2
Piano	100Hz	0	22.7	**77.3**	0	27.3	**72.7**
	150Hz	0	22.7	**77.3**	0	18.2	**81.8**
	200Hz	0	22.7	**77.3**	4.6	31.8	**63.6**
	250Hz	0	45.5	**54.5**	13.6	27.3	**59.1**
Marimba	100Hz	13.6	**45.4**	41	13.6	**45.5**	40.9
	150Hz	9.1	22.7	**68.2**	13.6	13.7	**72.7**
	200Hz	9.1	27.3	**63.6**	0	18.2	**81.8**
	250Hz	18.2	22.7	**59.1**	22.7	31.8	**45.5**
Saxophone	100Hz	36.4	**54.5**	9.1	36.4	**54.5**	9.1
	150Hz	**40.9**	**40.9**	18.2	40.9	**45.5**	13.6
	200Hz	9.1	**59.1**	31.8	9.1	**54.5**	36.4
	250Hz	40.9	**45.5**	13.6	22.7	**63.6**	13.7

4. CONCLUSION

Audio-visual mappings are very important for designing effective interactive musical interfaces and auditory-visual substitution systems. Current visual to auditory substitution systems generate many simple tones (sine waves) for an image using basic mappings such as vertical position of a pixel to pitch and pixel value to auditory loudness [9]. Real-time interpretation of visual scenes and identification of visual objects using a mixture of many sine waves are very difficult. Another strategy is to convert each visual object to a rich sound. Considering the apparent correspondence between timbre and shapes revealed by this study, conversion of a visual object to a sound might be achieved by mapping the visual shape of the object to a proper timbre i.e. an object with a jagged shape can be mapped to the windowed sound of Piano while an object with both sharp and rounded angles can be mapped to the windowed sound of Sax. Thus, complicated visual objects could be encoded in richer sounds which are much easier to interpret for the users.

5. ACKNOWLEDGMENT

We would like to thank all anonymous subjects, D. Bélisle for his useful comments on designing a protocol for the experiment, members of NECOTIS and the three anonymous reviewers for their constructive comments. This work was supported by grant ACELP-3IT from Université de Sherbrooke and grant FRNTQ from Québec government.

6. REFERENCES

[1] L. E. Marks, "On associations of light and sound: The mediation of brightness, pitch, and loudness," *The American journal of psychology*, vol. 87, no. 1–2, pp. 173–188, 1974.

[2] T. L. Hubbard, "Synesthesia-like mappings of lightness, pitch, and melodic interval," *The American journal of psychology*, vol. 109, no. 2, pp. 219–238, 1996.

[3] K. K. Evans and A. Treisman, "Natural cross-modal mappings between visual and auditory features," *Journal of vision*, vol. 10, no. 1, pp. 6.1–12, 2010.

[4] A. Abbado, "Perceptual correspondences of abstract animation and synthetic sound," *Leonardo. Supplemental Issue*, vol. 1, pp. 3–5, 1988.

[5] T. Grill and A. Flexer, "Visualization of perceptual qualities in textural sounds," in *Proc. of the Int. Computer Music Conf. (ICMC 2012)*, Ljubljana, Slovenia, 2012, pp. 589–596.

[6] K. Giannakis, "A comparative evaluation of auditory-visual mappings for sound visualisation," *Organised sound*, vol. 11, no. 03, pp. 297–307, 2006.

[7] V. S. Ramachandran and E. M. Hubbard, "Synaesthesia–a window into perception, thought and language," *Journal of Consciousness Studies*, vol. 8, no. 12, pp. 3–34, 2001.

[8] http://audio-visual-experiment.espaceweb.usherbrooke.ca.

[9] P. B. L. Meijer, "An experimental system for auditory image representations," *Biomedical Engineering, IEEE Transactions on*, vol. 39, no. 2, pp. 112 –121, 1992.